

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

1c978 U.S. PTO  
09/808407  
03/14/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2000年 3月14日

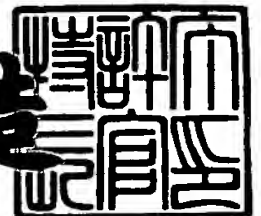
出願番号  
Application Number: 特願2000-070915

出願人  
Applicant(s): 日立ソフトウェアエンジニアリング株式会社

2001年 1月26日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2001-3002014

【書類名】 特許願

【整理番号】 11B114

【提出日】 平成12年 3月14日

【あて先】 特許庁長官 殿

【国際特許分類】 G01N 33/50

【発明の名称】 ハイブリダイゼーション実験の結果表示方法

【請求項の数】 7

【発明者】

    【住所又は居所】 神奈川県横浜市中区尾上町 6 丁目 8.1 番地 日立ソフト  
ウエアエンジニアリング株式会社内

    【氏名】 田村 卓郎

【発明者】

    【住所又は居所】 神奈川県横浜市中区尾上町 6 丁目 8 1 番地 日立ソフト  
ウエアエンジニアリング株式会社内

    【氏名】 山本 宣之

【発明者】

    【住所又は居所】 神奈川県横浜市中区尾上町 6 丁目 8 1 番地 日立ソフト  
ウエアエンジニアリング株式会社内

    【氏名】 鈴木 伸明

【発明者】

    【住所又は居所】 神奈川県横浜市中区尾上町 6 丁目 8 1 番地 日立ソフト  
ウエアエンジニアリング株式会社内

    【氏名】 水野 克也

【特許出願人】

    【識別番号】 000233055

    【氏名又は名称】 日立ソフトウエアエンジニアリング株式会社

【代理人】

    【識別番号】 100091096

    【弁理士】

【氏名又は名称】 平木 祐輔

【選任した代理人】

【識別番号】 100102576

【弁理士】

【氏名又は名称】 渡辺 敏章

【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9722155

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ハイブリダイゼーション実験の結果表示方法

【特許請求の範囲】

【請求項1】 バイオチップ上に固定した複数のプローブ生体高分子とサンプル生体高分子との間のハイブリダイゼーション実験の結果を表示する方法において、

ハイブリダイゼーション実験によって得られた各プローブ生体高分子におけるハイブリダイゼーション量の情報と共に各プローブ生体高分子間の塩基配列の類似度を表わす類似性スコアを表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項2】 請求項1記載のハイブリダイゼーション実験の結果表示方法において、類似性スコアの大小を色の濃淡で表現して表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項3】 請求項1記載のハイブリダイゼーション実験の結果表示方法において、類似性スコアの大小を色の濃淡で表現し、縦軸と横軸に対象プローブ生体高分子群を配してマトリックス状に並べて表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項4】 請求項1, 2又は3記載のハイブリダイゼーション実験の結果表示方法において、ハイブリダイゼーション量の情報はその大小を色の濃淡で表現して表示する、あるいは該当するプローブ生体高分子のスポットイメージによって表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項5】 請求項1～4のいずれか1項記載のハイブリダイゼーション実験の結果表示方法において、特定の1つのプローブ生体高分子と各プローブ生体高分子との間の類似性スコアの値でソートして、プローブ生体高分子のデータと、ハイブリダイゼーション量と、類似性スコアとを並べて表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項6】 請求項5記載のハイブリダイゼーション実験の結果表示方法において、複数のバイオチップに由来するハイブリダイゼーション量を並べて表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【請求項7】 請求項6記載のハイブリダイゼーション実験の結果表示方法において、対象プローブ生体高分子の、前記複数のバイオチップにおけるハイブリダイゼーション量の変化のプロファイルを統計的手法で解析してプローブ生体高分子を分類した結果と並べて表示することを特徴とするハイブリダイゼーション実験の結果表示方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、バイオチップを用いた配列が既知のプローブ生体高分子とサンプル生体高分子とのハイブリダイゼーション実験の結果を表示する表示方法に関する。

【0002】

【従来の技術】

サンプル中に、異なるボリュームで含まれる多種類の生体高分子例えばDNA配列を、一度に定量的に測定する技術として、バイオチップ（DNAマイクロアレイ）技術が開発されてきた。この技術の概要は、ペンシルバニア大学のVivian G. Cheung等によるレビューに纏められている（Vivian G. Cheung et al.: Making and reading microarrays, Nature Genetics Supplement, Vol.21, January 1999）。バイオチップ技術は、スライドガラス等の支持体上に固定された多種類のプローブ生体高分子（例えばDNA）に対して、標識を付けたサンプル中の多種類の生体高分子（例えばDNA配列）がハイブリダイゼーション反応により選択的に結合することを利用している。すなわち、プローブ生体高分子上にハイブリダイゼーション反応によってサンプル生体高分子を介して選択的に結合した標識の量により、特定の種類のサンプル生体高分子量を定量できる仕組みであり、非常に多くの種類のプローブ生体高分子を同一支持体上に固定することによって、一度に膨大な種類のサンプル生体高分子の定量を行うことを可能としている。

【0003】

ところで、2つのDNA配列がハイブリダイゼーション反応により結合するた

めには、2つの配列が相補配列、または相補に近い配列である必要がある。相補配列は結合エネルギーが高く一定の熱に対して安定であるが、この結合エネルギーは配列の長さや組成（GC含量）により値が異なる。また、たとえ2つの配列の間に相補で無い部分が存在しても、周辺に十分な長さの相補配列が存在すれば、十分な結合エネルギーを持ち得る。すなわち、良く似てはいるが異なる2種類のプローブDNAに対して、同一種類のサンプルDNAがハイブリダイゼーション反応により結合する可能性が存在する。ハイブリダイゼーション実験の実験条件により、この様な意図しないハイブリダイゼーション（ミスハイブリダイゼーション）が起こる可能性が増減することが知られている。

#### 【0004】

合成短鎖DNAをプローブDNAとして用いるバイオチップ（オリゴヌクレオチドアレイ）においては、各対象プローブDNAに類似した一連のDNA配列を合成し、これを比較対照用プローブDNAとすることにより、ハイブリダイゼーションにより結合したサンプルDNAが、対象プローブDNAがターゲットとしていた配列であることを検証する手法が知られており、これは、Affymetrix社のRobert J. Lipshutz等によるレビューに纏められている（Robert J. Lipshutz et al.: High density synthetic oligonucleotide arrays, Nature Genetics Supplement, Vol.21, January 1999）。

しかしながら、中、長鎖DNA、たとえばcDNAをプローブ生体高分子として用いるバイオチップにおいては、DNA配列データを用いた効果的なハイブリダイゼーション反応検証の手法は報告されていない。

#### 【0005】

一方、異なる2つのDNA配列間で、最も相同性の高い領域を探索する手法として、SMITH-WATERMAN法（Smith, T. F. and Waterman, M. S.: J. Mol. Biol. 147, 195-197, 1981）が知られている。また、多種類のDNA配列（ターゲット）中より、目的のDNA配列（キー配列）と相同性の高いターゲットを高速に探索する方法として、BLAST法（Altschul et al., Nucleic Acids Res., 25, 3389-3402, 1997）が知られている。その他、同様の目的のアルゴリズムが多数、開発されている。これらの手法においては、探索の結果として、2つのDNA配列間

における相同性の高い領域について、探索に用いたスコアに基づく「相同性スコア」や、その範囲において相補であるDNAの割合に基づく「一致率」を指標として相同性の度合いを表現する（本明細書においては、相同性の度合いを表すこれらの指標を総称して「類似性スコア」と称する）。

#### 【0006】

また、バイオチップ実験情報解析では、対象プローブDNA群を、複数のバイオチップにおけるハイブリダイゼーション量の変化によって統計的に分類（クラスタリング）する手法が広く用いられている。スタンフォード大学のP.Brownらのグループによるイースト菌の発現解析（Michel B.Eisen et al.: Cluster analysis and display of genome-wide expression patterns: Proc. Natl. Acad. Sci. (1998) Dec 8, 95 (25), 14863-8）では、細胞のステージ毎（時系列）にサンプルDNAを調整し、別々のバイオチップとハイブリダイズさせ、各ステージのサンプルDNAにおいて存在するDNA配列の種類と量を測定し、各DNA配列のステージ毎の変化量によって、DNA配列（プローブDNA）をクラスタリングしている。結果の表示においては、クラスタリングによって求められた、DNA配列におけるクラスタ併合の順番とクラスタ間距離を示す樹状図、DNA配列の情報（名称、定義情報等）、そして各バイオチップにおける各DNA配列のハイブリダイゼーション量をハイブリ量パターンで表示している。

#### 【0007】

##### 【発明が解決しようとする課題】

中、長鎖DNAをプローブ生体高分子とするバイオチップにおいては、現在のところハイブリダイゼーション反応によりプローブ生体高分子が目的とするサンプル生体高分子と正確に結合できているかどうかを判定するための簡便な方法がなく、そのような方法の開発が待たれている。

本発明は、このような要請に応えるべくなされたもので、バイオチップを用いたハイブリダイゼーション実験の精度に関する情報を視覚的に分かりやすく表示する方法を提供することを目的とする。

#### 【0008】

##### 【課題を解決するための手段】

本発明では、前記目的を達成すべく、対象プローブ生体高分子の塩基配列より類似性スコアを計算し、これを濃淡のある矩形等のパターン（類似性パターン）で表現して、プローブ生体高分子のデータ及びハイブリダイゼーション量のデータと並べて比較表示することによって、目的プローブ生体高分子と類似の塩基配列を持つプローブ生体高分子が、類似のハイブリダイゼーション量であるかないかを視覚的に確認可能とすることで、意図しないハイブリダイゼーション反応の有無を容易に知ることを可能とする。また、複数のバイオチップにおける対象プローブ生体高分子のハイブリダイゼーション量情報を同時に表示することによって、ハイブリダイゼーション反応に不都合のあったバイオチップが無いことを確認可能とする。この際、類似性パターンを、対象プローブ生体高分子を縦軸、横軸に配してマトリックス状に表示（類似性パターンマトリックス）することで、より直感的な確認が可能な表示を提供する。

#### 【0009】

加えて、複数のバイオチップを対象としたクラスタ解析の結果と、類似性パターンマトリックスを並べて表示することにより、クラスタ解析が、塩基配列の物理特性でなく、生物学的特性により分離されていることの確認を可能とする。

すなわち、本発明によるハイブリダイゼーション実験の結果表示方法は、バイオチップ上に固定した複数のプローブ生体高分子とサンプル生体高分子との間のハイブリダイゼーション実験の結果を表示する方法において、ハイブリダイゼーション実験によって得られた各プローブ生体高分子におけるハイブリダイゼーション量の情報と共に各プローブ生体高分子間の塩基配列の類似度を表わす類似性スコアを表示することを特徴とする。

#### 【0010】

類似性スコアの大小は、色の濃淡で表現して表示することができる。また、類似性スコアの大小を色の濃淡で表現し、縦軸と横軸に対象プローブ生体高分子群を配してマトリックス状に並べて表示することもできる。

ハイブリダイゼーション量の情報はその大小を色の濃淡で表現して表示したり、あるいは該当するプローブ生体高分子のスポットイメージによって表示することができる。



## 【 0 0 1 1 】

特定の1つのプローブ生体高分子と各プローブ生体高分子との間の類似性スコアの値でソートして、プローブ生体高分子のデータ（名称、定義情報等）と、ハイブリダイゼーション量と、類似性スコアとを並べて表示することは有効な表示方法である。複数のバイオチップに由来するハイブリダイゼーション量を並べて表示することも有効である。

また、対象プローブ生体高分子の、複数のバイオチップにおけるハイブリダイゼーション量の変化のプロファイルを統計的手法（クラスタ解析等）で解析してプローブ生体高分子を分類した結果と並べて表示することもできる。

## 【 0 0 1 2 】

プローブ生体高分子の塩基配列情報より類似性スコアを計算し、濃淡のある矩形等のパターンでこれを表わす。複数の対象プローブ生体高分子間ですべての類似性スコアを計算し、矩形をマトリックス状に並べることで、類似性スコアマトリックスとし、各対象プローブのハイブリダイゼーション量情報に並べて表示する。マトリックスとハイブリダイゼーション量情報の表示順を、目的とする1つのプローブ生体高分子と、他のプローブ生体高分子との間の類似性スコアでソートすると、目的のプローブと類似度の高い配列でソートされるが、その際、目的プローブの周辺のハイブリダイゼーション量情報を観察することにより、意図しないハイブリダイゼーション反応が起こっていないか、確認することが可能となる。類似性スコアマトリックスと同時に表示する情報を選択することで、より広範囲なハイブリダイゼーションの精度の検証も可能となる。

## 【 0 0 1 3 】

## 【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。以下では、プローブ生体高分子及びサンプル生体高分子がDNAである場合を例にとって本発明を説明する。しかし、本発明はプローブDNAとサンプルDNAのハイブリダイゼーション実験のみでなく、プローブDNAとサンプルDNA結合蛋白質、モノクローナル抗体とサンプル蛋白質など他の生体高分子のハイブリダイゼーション実験一般に適用できることは勿論である。

## 【0014】

図1は、本発明のシステム構成図である。このシステムは、バイオチップデータ（名称、定義情報、実験情報等）、バイオチップ上の各プローブDNAのハイブリダイゼーション量データ（ハイブリダイゼーション量、スポットイメージ等）、プローブDNA情報（名称、定義情報、DNA配列等）、その他のバイオチップデータを格納したバイオチップデータ格納部100、そのデータを視覚化して表示するための表示装置101、本システムへの値の入力や選択の操作を行うためのキーボード102、マウス等の入力装置103、DNA配列の類似性スコアの計算、類似性パターンの生成・ソート等の処理を行う情報処理部104から構成される。

## 【0015】

図2は、本発明による配列類似性パターン表示の一例の概略処理フローを示した図である。この処理フローにしたがって、順次説明する。

まず、ステップ200において、調査対象とする複数のプローブDNA（対象プローブDNA群）を指定する。次に、ステップ201において、対象プローブDNA群に含まれる各対象プローブDNAの情報を取得する。すなわち、ステップ200で選択された対象プローブDNAについて、それぞれ、プローブDNAデータ（プローブDNA名、プローブ定義情報、プローブDNA配列等）をバイオチップデータ格納部100より取得する。次に、ステップ202において、対象プローブDNA群に含まれる2つのプローブDNA間の類似性スコアを、全ての組み合わせにおいて計算する。異なる2つのプローブDNA間のみでなく、同一プローブDNA間における類似性スコアも計算する。

## 【0016】

次に、ステップ203に進み、調査対象とするバイオチップ（対象バイオチップ）を指定する。対象バイオチップの選択は単独、複数の何れでも可能とする。次に、ステップ204において、各対象バイオチップにおける各対象プローブDNAのハイブリダイゼーション量データを取得する。すなわち、ステップ200で選択された対象プローブDNAとステップ203で選択された対象バイオチップにより、対象バイオチップ上にある対象プローブDNAのハイブリダイゼーシ

ョン量のデータを取得する。

【0017】

次に、ステップ205に進み、対象プローブDNA群より調査において着目するプローブDNA（目的プローブDNA）を1つ選択し、各対象プローブDNAを目的プローブDNAとの類似性スコアによりソートする。あるいは、ステップ206に進み、クラスタ解析等により、対象バイオチップ上の対象プローブDNAのハイブリダイゼーション量をデータ処理し、対象プローブDNAを分類、ソートする。最後にステップ207に進み、ステップ205もしくはステップ206でソートした順番で、対象プローブDNA毎に以上の処理結果を表示する。ステップ206経由の場合には、データ処理による付加情報が存在する場合は加えて表示する。

以下に、図3～図18に示す例を用いて図2のフローチャートに示した各ステップの処理の内容を詳細に説明する。

【0018】

図3は、バイオチップデータ格納部100に格納されているデータ構造の例を示す図である。バイオチップテーブル300には、バイオチップID、バイオチップ名、バイオチップ定義情報、及びバイオチップ実験情報が格納されている。プローブDNAテーブル301には、プローブDNAのID、プローブDNA名、プローブDNA定義情報、プローブDNA配列が格納されている。また、ハイブリダイゼーション量テーブル302には、バイオチップID、プローブDNAのID、ハイブリダイゼーション量、及びスポットイメージが格納されている。

【0019】

図4はバイオチップテーブル300に格納されているバイオチップデータの例、図5はプローブDNAテーブル301に格納されているプローブDNAデータの例、図6はハイブリダイゼーション量テーブル302に格納されているハイブリダイゼーション量データの例をそれぞれ示している。

【0020】

図2のステップ200では、バイオチップデータ格納部100に格納されたプローブDNAデータ301より対象プローブDNAを選択する。ここでは、プロ

ーブDNAデータ301において、プローブDNA IDが1～5のプローブDNAを選択し、対象プローブDNA群とした例によって説明する。ステップ201では、ステップ200で選択されたプローブDNA IDが1～5の各プローブDNAについて、プローブDNAデータ301よりプローブDNAデータ（プローブDNA名、プローブDNA定義情報、プローブDNA配列等）を取得する。

#### 【0021】

図7に、ステップ202で計算した類似性スコア計算例を示す。この類似性スコアは、ステップ201で取得したプローブDNA IDが1～5のプローブDNAのDNA配列を用いて、SMITH-WATERMAN法により相同性スコア（類似性スコア）を求め、マトリックス状に表示したものである。

#### 【0022】

ステップ203では、図4に示したバイオチップデータ300より対象バイオチップを選択する。例えば、バイオチップデータ300からバイオチップIDが1のバイオチップを選択し対象バイオチップとする場合、また、バイオチップIDが1～3のバイオチップを選択し、対象バイオチップとする場合などがあげられる。ステップ204では、例えば図6に示すハイブリダイゼーション量データ302a, 302b, ……より、対象プローブDNAの持つプローブDNA IDと、対象バイオチップの持つバイオチップIDにより選択されるデータレコードを選択し、各レコードのハイブリダイゼーション量データ（ハイブリダイゼーション量、スポットイメージ等）を取得する。

#### 【0023】

表示内容において、ステップ202において計算した類似性スコアは、例えば類似性スコアの値の大小を色の濃淡で表わした矩形パターン（類似性パターン）として表示を行う。すなわち、図8に示すように、計算された類似性スコアの最小値602と最大値601の間で色勾配603を用意し、類似性スコア値に対応する色を決定し、類似性パターン604に変換する。

#### 【0024】

具体例として、図9に示すように、プローブDNAデータ301のプローブD

NA ID: 1～5について、プローブDNA ID: 1のプローブDNAとの間の類似性スコア605を変換すると、類似性パターン606となる。また、この類似性パターンを、図7に図示した類似性スコア計算例と同様にマトリックス状に配列することで、類似性パターンマトリックスが表示できる。図10に類似性パターンマトリックスの例を示す。

#### 【0025】

また、表示内容において、ステップ204で取得したハイブリダイゼーション量は、例えばハイブリダイゼーション量の値の大小を色の濃淡で表わした矩形パターン（ハイブリ量パターン）として表示を行う。すなわち、図11に示すように、ハイブリダイゼーション量は、特定の最小値702と特定の最大値701の間で色勾配703を用意し、ハイブリダイゼーション量の値に対応する色を決定し、ハイブリ量パターン704に変換する。

#### 【0026】

例えば、図12に示すように、図6に示したハイブリダイゼーション量データ302aのプローブDNA ID: 1～5について、ハイブリダイゼーション量705を変換すると、ハイブリ量パターン706となる。本例においては、特定の最小値702及び特定の最大値701は、一般的なバイオチップ測定装置の測定レンジが2バイトオーダーであることから、そのレンジにおいて表現可能な値である0～65535を値として用いたが、実際には、処理対象となるハイブリダイゼーション量の範囲を反映することにより、より視覚的に差異を判別可能なハイブリ量パターンを生成することが可能となる。

#### 【0027】

図13～図18に処理結果表示の例を示す。以下の例においては、対象プローブDNA毎に1行にデータを表示し、対象プローブDNA間の情報を容易に閲覧可能としている。対象プローブ毎に表示するデータとしては、プローブDNAデータ806、ハイブリダイゼーション量データ807と、類似性パターン808がある。

#### 【0028】

図13は、図5に示したプローブDNAデータ301におけるProbe3を目的

プローブDNAとした場合の表示例である。この表示801においては、プローブDNAデータ806として、プローブDNA名、プローブDNA定義情報を、ハイブリダイゼーション量データ807として単一の対象バイオチップ上のスポットイメージを、目的プローブDNAに対する類似性スコアを変換した類似性パターン808と並べて表示している。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。この表示801から、対象プローブ内で、目的プローブProbe3とDNA配列の類似度が高いプローブとしてはProbe5が挙げられるが、Probe1のハイブリダイゼーション量に比較してProbe5のハイブリダイゼーション量は少なく、目的プローブProbe3に対応するサンプルDNAが、Probe5をはじめとする他の対象プローブにミスハイブリしている可能性が小さいことが読みとれる。

## 【0029】

図14は、図13の変形例である。この表示802においては、プローブDNAデータ806として、プローブDNA名、プローブDNA定義情報を、ハイブリダイゼーション量データ807として単一の対象バイオチップ上のハイブリ量パターンを、目的プローブDNAに対する類似性スコアを変換した類似性パターン808と並べて表示している。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。

## 【0030】

図15は、同様に、図5に示したプローブDNAデータ301におけるProbe3を目的プローブDNAとした場合の表示例である。この表示803においては、プローブDNAデータ806として、プローブDNA名、プローブDNA定義情報を、ハイブリダイゼーション量データ807として複数の対象バイオチップ上のハイブリ量パターンを、目的プローブDNAに対する類似性スコアを変換した類似性パターン808と並べて表示している。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。

## 【0031】

図16は、図15に示した表示例803と同様の表示において、図5に用意し

たプローブDNAデータ301において、プローブDNA ID: 1 (プローブ名称: Probe 1) を目的プローブDNAとして表示した例である。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。ここにおいて、ハイブリダイゼーション量データ807中の枠809で囲んだ部分から、Chip2におけるProbe 1とProbe 2が非常に近いハイブリダイゼーション量であることが、また類似性パターン808中の枠810で囲んだ部分より、Probe 1とProbe 2のプローブDNA配列が類似度の高い配列であることが分かる。この様な組み合わせがミスハイブリの可能性のある個所として容易に識別できる。

## 【0032】

図17は、類似性パターンマトリックスによる表示例を示している。プローブDNAデータ806として、プローブDNA名、プローブDNA定義情報を、ハイブリダイゼーション量データ807として複数の対象バイオチップ上のハイブリ量パターンを表示し、加えて、類似性パターンを対象プローブDNA間すべての組み合わせにおいてマトリックス状に表示した類似性パターンマトリックス901を並べて表示している。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。この表示方式においては、目的プローブDNA以外のプローブDNAについても、対象プローブDNA全体に渡るDNA配列が類似のプローブDNA間におけるハイブリダイゼーション量データとの関連を概観することが可能である。

## 【0033】

図18は、類似性パターンマトリックスとクラスタ解析結果を合わせて表示した表示例である。図2のステップ206において、対象バイオチップ上の対象プローブDNAのハイブリダイゼーション量をもとに、各バイオチップにおけるハイブリダイゼーション量変化を対象としてプローブDNAを、また、各プローブDNAにおけるハイブリダイゼーション量の変化を対象としてバイオチップをクラスタ処理し、結果をそれぞれ1001、1002として樹状表示した。

## 【0034】

図示するように、プローブDNAでクラスタ処理した結果をもとに対象プロ

ブDNAをソートし、対象プローブDNA毎に、プローブDNAデータ806としてプローブDNA名とプローブDNA定義情報を、ハイブリダイゼーション量データ807として複数の対象バイオチップ上のハイブリ量パターンを表示し、加えて、類似性パターンを対象プローブDNA間すべての組み合わせにおいてマトリックス状に表示した類似性パターンマトリックス901を並べて表示している。ハイブリダイゼーション量データ807の対象バイオチップに関しては、別途設けた表示部805にその名称を表示している。

#### 【0035】

この表示方式においては、対象バイオチップ全般に渡って、ハイブリダイゼーション量データの解析の結果が、ミスハイブリによりプローブDNAの物理的特徴（プローブDNAのDNA配列）を反映したものであるか、もしくは、サンプルの生物学的な特徴（サンプル中のDNA配列の存在状況）を反映したものであるかを概観することができる。図18の表示例では、類似性パターンマトリックス901の表示よりProbe1とProbe2が非常に類似度の高いDNA配列であることがわかるが、樹状図1001によると、比較的異なる性質（Probe1はProbe4とより近い性質である）と分類されていることから、解析結果がプローブDNAの物理的な特徴よりも、生物学的な特徴を反映していることが可能である。

#### 【0036】

以上のように、図2に示した処理フローに従うことによって、図13～図18に示したような表示が可能となる。また、処理結果表示の後、ステップ208により、目的プローブDNAを変更し、異なる目的プローブDNAに対する同様の表示を行い、ハイブリダイゼーション精度の検証を行うことが可能である。

#### 【0037】

##### 【発明の効果】

本発明によると、対象プローブDNAのDNA配列を利用して、バイオチップ技術におけるハイブリダイゼーション実験の精度の検証を可能とする簡便な表示方式を提供できる。

##### 【図面の簡単な説明】



【図 1】

本発明のシステム構成図。

【図 2】

本発明による配列類似性パターン表示の一例の概略処理フローを示した図。

【図 3】

バイオチップデータ格納部に格納されているデータ構造の例を示す図。

【図 4】

バイオチップテーブルに格納されているバイオチップデータの例を示す図。

【図 5】

プローブ DNA テーブルに格納されているプローブ DNA データの例を示す図。

【図 6】

ハイブリダイゼーション量テーブルに格納されているハイブリダイゼーション量データの例を示す図。

【図 7】

類似性スコアの計算例を示す図。

【図 8】

類似性パターンの説明図。

【図 9】

類似性パターン作成の例を説明する図。

【図 10】

類似性パターンマトリックスの説明図。

【図 11】

ハイブリ量パターン生成の例を示す図。

【図 12】

ハイブリ量パターン変換の例を示す図。

【図 13】

類似性パターン表示方式の一例を示す図。

【図 14】

類似性パターン表示方式の他の例を示す図。

【図 1 5】

類似性パターン表示方式の他の例を示す図。

【図 1 6】

類似性パターン表示方式の他の例を示す図。

【図 1 7】

類似性パターンマトリックス表示方式の一例を示す図。

【図 1 8】

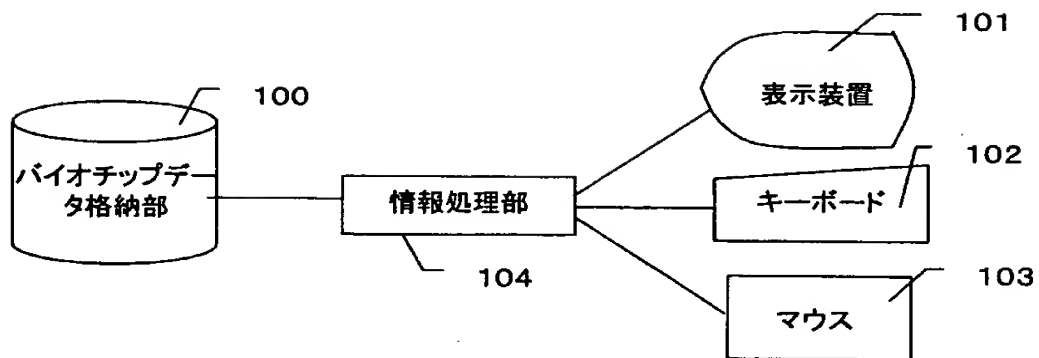
クラスタ解析結果と類似性パターンマトリックスの表示例を示す図。

【符号の説明】

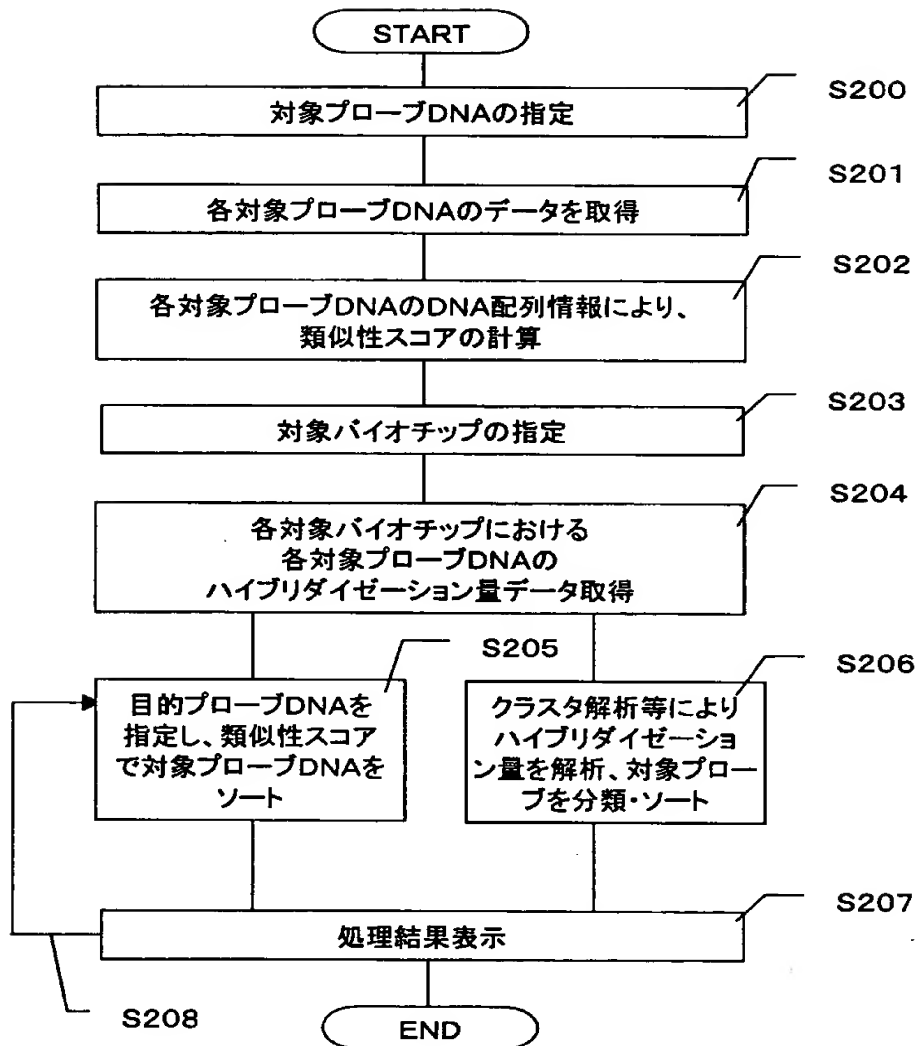
1 0 0 … バイオチップデータ格納部、1 0 1 … 表示装置、1 0 2 … キーボード、1 0 3 … マウス、1 0 4 … 情報処理部、3 0 0 … バイオチップテーブル（バイオチップデータ）、3 0 1 … プローブDNAテーブル（プローブDNAデータ）、3 0 2 … ハイブリダイゼーション量テーブル（ハイブリダイゼーション量データ）、6 0 3 … 色勾配、6 0 4 … 類似性パターン、6 0 5 … 類似性スコア、6 0 6 … 類似性パターン、7 0 3 … 色勾配、7 0 4 … ハイブリ量パターン、7 0 5 … ハイブリダイゼーション量、7 0 6 … ハイブリ量パターン、8 0 6 … プローブDNAデータ、8 0 7 … ハイブリダイゼーション量データ、8 0 8 … 類似性パターン、9 0 1 … 類似性パターンマトリックス、1 0 0 1 … 樹状図

【書類名】 図面

【図 1】

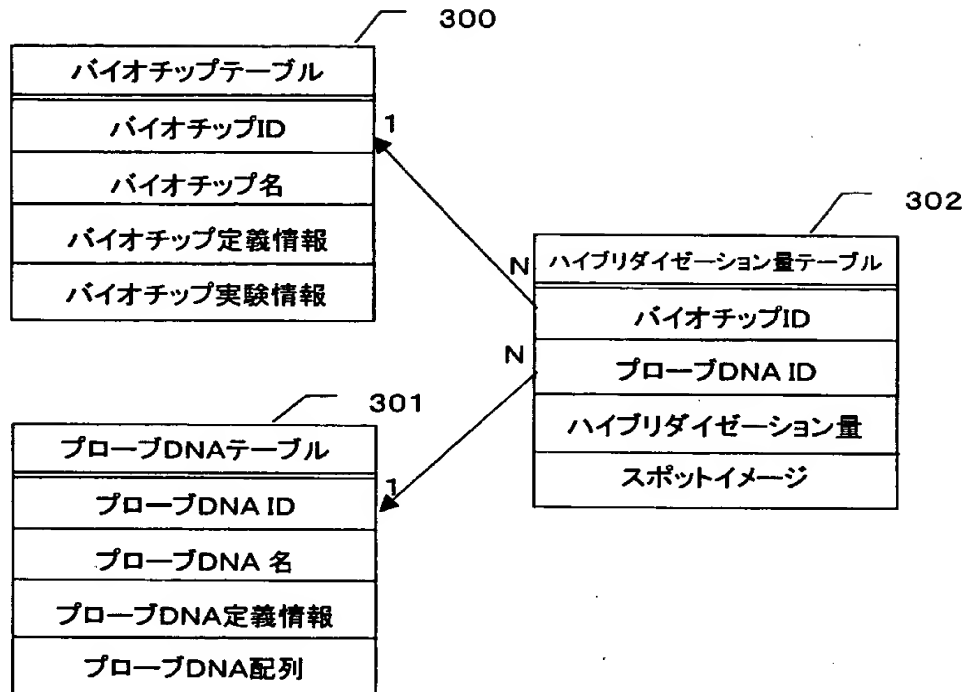


【図 2】



【図 3】

データ構造の例



【図 4】

300 バイオチップデータの例

バイオチップID	バイオチップ名	バイオチップ定義情報	バイオチップ実験情報
1	Chip 1	Definition about Chip 1	Sample from the Stage A (1D)
2	Chip 2	Definition about Chip 2	Sample from the Stage B (3D)
3	Chip 3	Definition about Probe 3	Sample from the Stage C (1W)
⋮	⋮	⋮	
n	Chip n	Definition about Chip n	not used.

【図 5】

301 プローブDNAデータの例

プローブ DNA ID	プローブ DNA名	プローブDNA定義情報	プローブDNA配列
1	Probe1	Definition about Probe 1	ACGGGACGTTCCCTCGGAGG
2	Probe2	Definition about Probe 2	ACGGGACGTTCCCCCGGAGG
3	Probe3	Definition about Probe 3	TTTGGACGTTCCAATAGAGG
4	Probe4	Definition about Probe 4	TTTGGACGTTCCAATAGGGG
5	Probe5	Definition about Probe 5	TTGGGACGTTCCAATAGGAG
:	:	:	:
m	Probe m	Definition about Probe n	TTGGGACGTTCCAATAGGAG

【図 6】

ハイブリダイゼーション量データの例

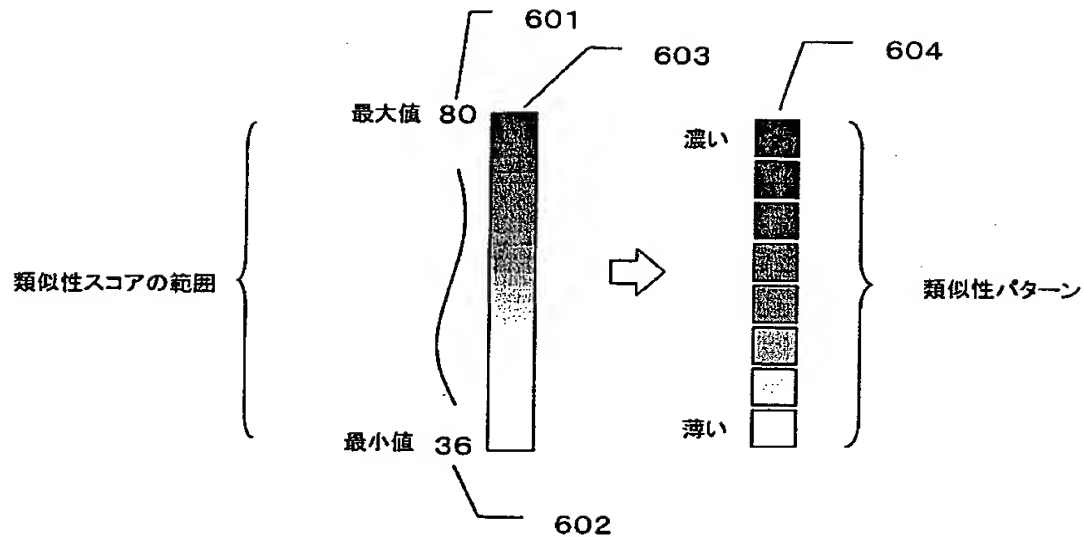
302a				302b				302c			
バイオ チップID	プローブ DNA ID	ハイブリダイ ゼーション量	スポット イメージ	バイオ チップID	プローブ DNA ID	ハイブリダイ ゼーション量	スポット イメージ	バイオ チップID	プローブ DNA ID	ハイブリダイ ゼーション量	スポット イメージ
1	1	31,916		2	1	43,416		3	1	21,916	
1	2	224		2	2	41,224		3	2	11,224	
1	3	41,231		2	3	2,107		3	3	10,007	
1	4	24,107		2	4	25,307		3	4	21,107	
1	5	9,607		2	5	1,807		3	5	1,507	
:	:	:		:	:	:		:	:	:	
1	m	9,753		2	m	853		3	m	5,753	

【図 7】

類似性スコア計算例

Key Target	Probe1	Probe2	Probe3	Probe4	Probe5
Probe1	80	76	40	56	44
Probe2	76	80	44	52	44
Probe3	40	44	80	36	52
Probe4	56	52	36	80	48
Probe5	44	44	52	48	80

【図 8】

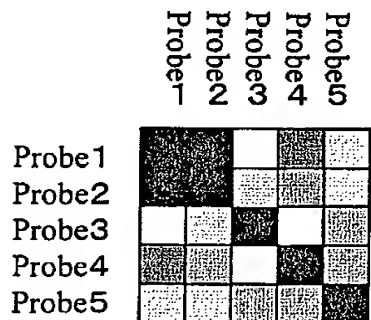


【図 9】

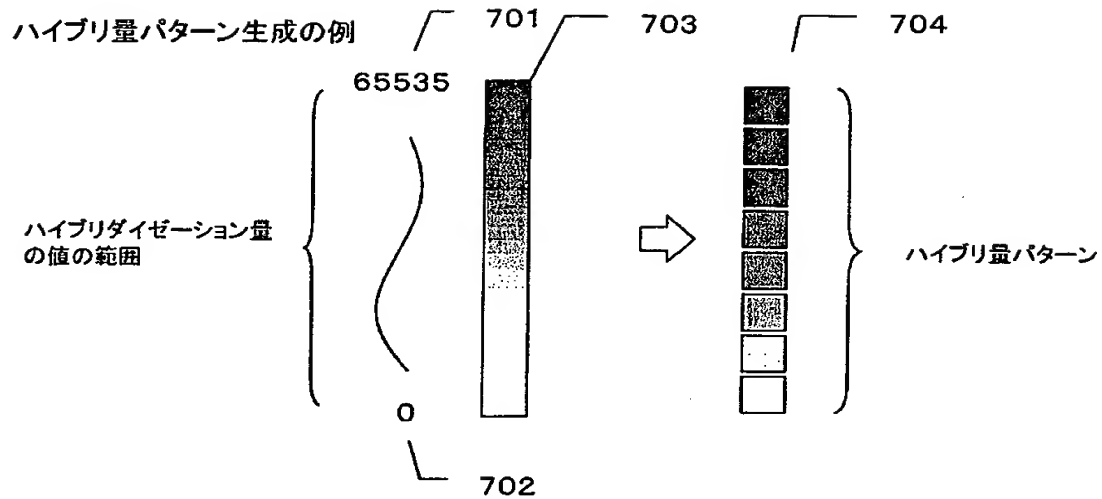
301				605	606
プローブ DNA ID	プローブ DNA名	プローブDNA定義情報	プローブDNA配列	類似性 スコア	類似性 パターン
1	Probe1	Definition about Probe 1	ACGGGACGTTCCCTCGGAGG	80	
2	Probe2	Definition about Probe 2	ACGGGACGTTCCCCCGGAGG	76	
3	Probe3	Definition about Probe 3	TTTGGACGTTCCAATAGAGG	40	
4	Probe4	Definition about Probe 4	TTTGGACGTTCCAATAGGGG	56	
5	Probe5	Definition about Probe 5	TTGGGACGTTCCAATAGGAG	44	
:	:	:	:		
m	Probe m	Definition about Probe n	TTGGGACGTTCCAATAG...	nn	

\* 類似性スコアは、プローブDNA ID 1に対する各プローブDNAの類似性スコア

【図 10】

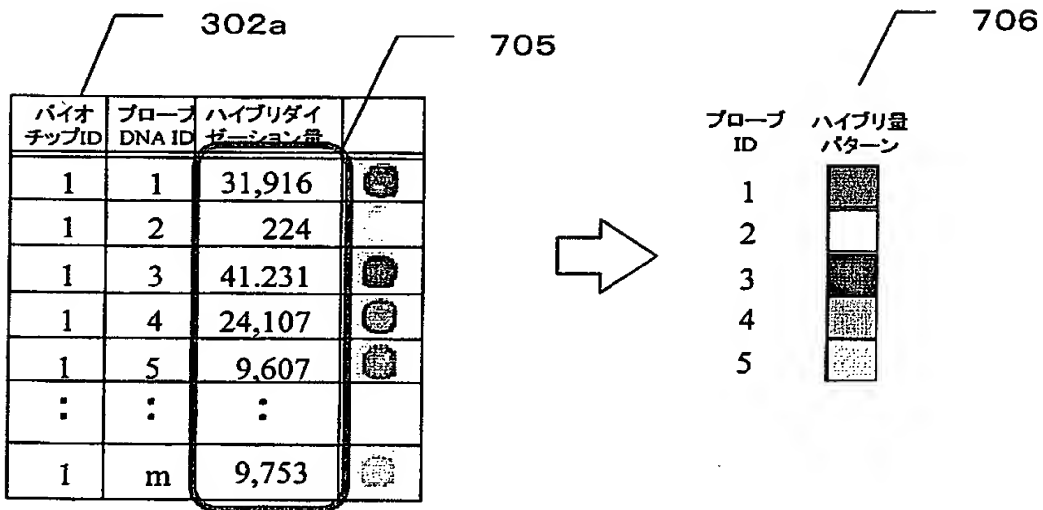


【図 1 1】

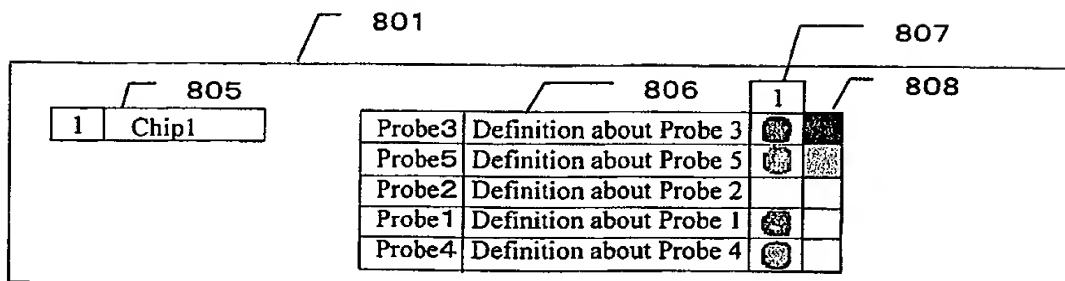


【図 1 2】

ハイブリ量パターン変換の例

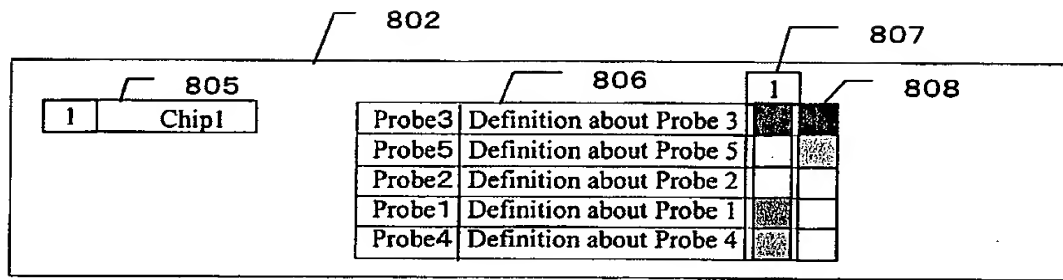


【図 1 3】

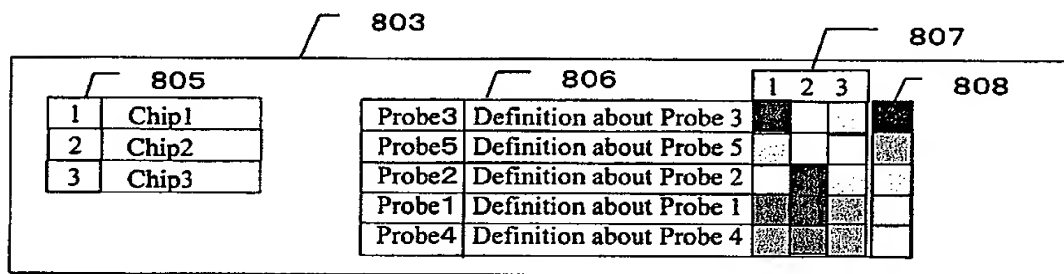




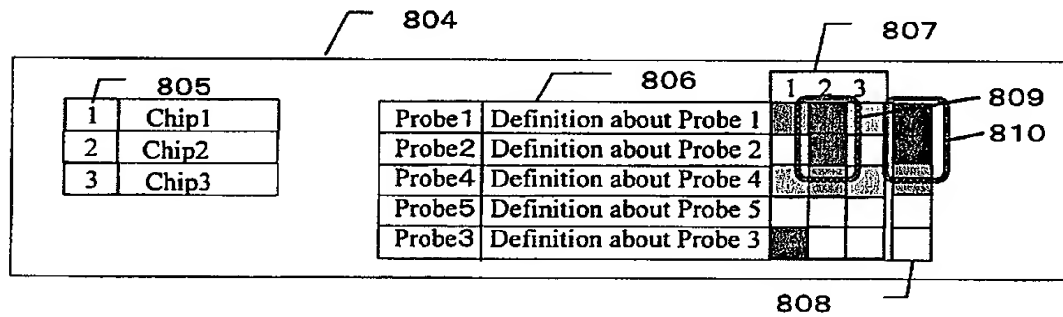
【図 1 4】

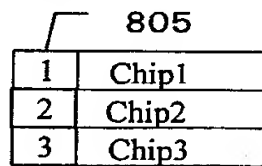
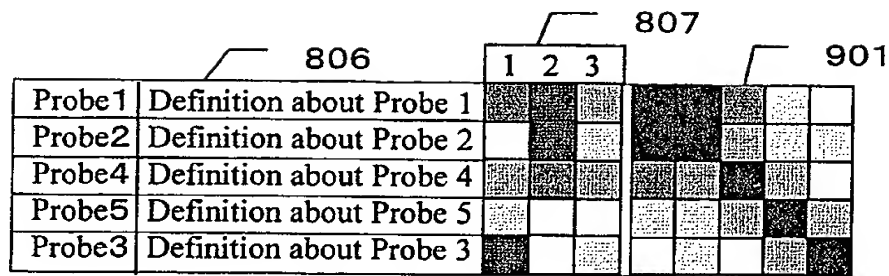


【図 1 5】

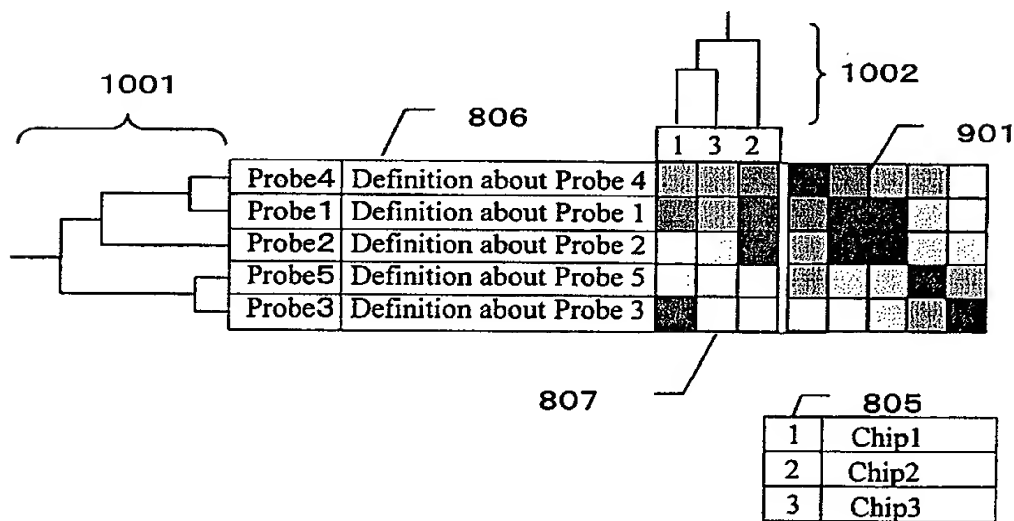


【図 1 6】





【图 18】



【書類名】 要約書

【要約】

【課題】 ハイブリダイゼーション実験の精度の検証を可能とする。

【解決手段】 バイオチップ上に固定した複数のプローブ生体高分子とサンプル生体高分子との間のハイブリダイゼーション実験の結果を表示するとき、ハイブリダイゼーション実験によって得られた各プローブ生体高分子におけるハイブリダイゼーション量の情報 8 0 7 と共に各プローブ生体高分子間の塩基配列の類似度を表わす類似性スコア 9 0 1 を表示する。

【選択図】 図 1 8

出 願 人 履 歴 情 報

識別番号 [000233055]

1. 変更年月日 1990年 8月 7日  
[変更理由] 新規登録  
住 所 神奈川県横浜市中区尾上町6丁目81番地  
氏 名 日立ソフトウェアエンジニアリング株式会社